DOCUMENT RESUME

ED 373 426 EA 026 035

AUTHOR Elliott, Emerson J.
TITLE Research and Reform: Stories of Structure, Strategy and Suspense.
PUB DATE Apr 94
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 4-8, 1994). Appendices contain smudged print and may not reproduce well.
PUB TYPE Speeches/Conference Papers (150) -- Viewpoints (Opinion/Position Papers, Essays, etc.) (120)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Blue Ribbon Commissions; *Educational Assessment; *Educational Policy; *Educational Research; Elementary Secondary Education; Federal Programs; Policy Formation; *Program Evaluation; Research Committees; *Theory Practice Relationship

ABSTRACT

This paper offers examples of the ways in which research influences national policies related to assessment and program evaluation. It describes an ongoing evaluation being conducted by the National Academy of Education (NAE) of the "Trial State Assessment," a piece of the National Assessment of Educational Progress (NAEP). At this interim stage, the NAE study has been well received by Congress and other audiences because the NAE panel: (1) covered the topics requested by Congress; (2) did its homework; (3) comported itself impartially; (4) provided clearly written results; (5) observed Washington etiquette regarding submission; and (6) briefed Congress personally. Any external evaluation group must guarantee impartial research, overcome Congressional distrust of researchers, resist pressure from contending parties, move from broad legislative objectives to specific researchable problems, and specify deadlines and formats for clear reporting. A conclusion is that researchers can strongly influence national policy when they have evidence for their conclusions; their work deals with problems faced by policymakers; and they clearly state their conclusions. Appendices contain pertinent legislation from the General Education Provisions Act, a list of studies completed by the NAE evaluation panel, and topics of NAE research. (LMI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. Elliott
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED 373 426

Remarks prepared by Emerson J. Elliott
for AERA session 1.11, Monday April 4, 1994
"Research and Reform: Stories of Structure, Strategy and Suspense"

INTRODUCTION

My role as a member of this panel is to provide examples of "How research influences national policies related to assessment and program evaluation."

Frequently we hear words of despair about the effect of the work of researchers on public policy decisions:

- o researchers don't understand our problems
- o researchers write technical jargon
- o researchers deal in theories rather than issues

All true.

Also true, however, is that the work of researchers is USED in decisionmaking, is frequently SOUGHT OUT.

EXAMPLES OF RESEARCH THAT INFLUENCED NATIONAL POLICIES

Let me mention some examples of panels and studies that have been comprised of, or have drawn on, researchers and their work in the areas of assessment and program evaluation.

- o The 1983 National Commission on Excellence in Education, which produced its report, A Nation at Risk, commissioned 41 papers, made extensive use of statistical data, and drew on in-depth analyses, to drive home its points. The report was a major factor in precipitating State legislated education reforms during the 1980s--reforms that failed to achieve the results policymakers intended and that set the stage for the more recent adoption of National Education Goals.

- o The 1989 National Assessment of Vocational Education contracted for 30 inter-related research studies. It described the implementation of Federal vocational education laws and concluded that the statutory provisions were too weak to accomplish the Congressional goals. The Assessment provided information, available for the first time from NCES studies, that showed most high school students take at least one or two vocational education

EA026036

BEST COPY AVAILABLE

courses, although the traditional concept of vocational "programs" was not commonly applied and not useful for understanding the effects of the law. As recommended by the Assessment, Congress dropped numerous categorical set asides in the Act, provided a floor on the average grant size, distinguished secondary from post-secondary activities, and encouraged a strong link between academic and work-skill content for vocational students.

- The 1987 Alexander-James study group on the National Assessment of Educational Progress commissioned 46 Research papers, included an independent critique from members of the National Academy of Education, and convened nine subgroups involving 64 experts, mostly researchers, on topics ranging from cognitive skills assessments and reading assessment to design and structure of NAEP and costs. The Panel's recommendations provided the basis for legislative changes in 1988 that authorized a State component for NAEP and created the National Assessment Governing Board.

- The National Education Goals Panel has created "resource groups," "technical planning groups" and has drawn on expert consultants for each of the Goals. The most recent listing included 146 individuals, a heavy proportion of them academic researchers, involved in these capacities. Another 91 individuals assisted the Panel in acquiring data for use in the annual report on U. S. progress toward the goals. As you know, the topics here include readiness of children for school, school completion, student achievement, adult literacy and college achievement, and safe schools--all now enacted into the GOALS 2000 legislation that President Clinton has just signed into law.

- The Federal Government's primary grant of assistance for education of disadvantaged students, Chapter I, has seen extensively evaluated over its nearly three decade existence. Within the last year we have witnessed the completion of (a) the National Assessment of Chapter I, mandated by Congress four years ago, of (b) an Independent Review Panel for the National Assessment, also mandated by law, of (c) an Independent Commission on Chapter I (Chaired by David Hornbeck and funded by the MacArther Foundation) as well as (d) a RAND study entitled "Federal Policy Options for Improving the Education of low-income students." These studies and panels--each drawing on members of the research community for advice or conduct of empirical work or analysis--share a number of common perspectives. They assert that the learning goals for

low-income or disadvantaged children must be the same as those for all our children. They call for "school-wide" approaches to instruction, rather than pulling children out of class. They propose that professional development, instruction, and assessments be linked to curricular goals. And they call for assessments more closely related to the actual knowledge and skills that students are expected to master.

Well, there could be many more such examples (see Note A) where the work of researchers related to assessments and program evaluations have influenced Federal policies, but I think these will serve to demonstrate that such work is both sought out and used in making national policy.

A CASE STUDY--EVALUATION OF THE TRIAL STATE ASSESSMENT

Let me make this more concrete through a case study. The example is an on-going evaluation by the National Academy of Education of the "Trial State Assessment," a piece of the National Assessment of Educational Progress (NAEP).

The setting for this case study begins with recommendations for improvements in the NAEP made by the Alexander-James study group in March of 1987, after the House of Representatives had already completed its work on the legislative authorization that would ordinarily have included NAEP. The House, however, had no proposed modifications for the NAEP legal authority that year.

The Alexander-James recommendations, formulated into a legislative proposal by the Department of Education under then Secretary Bennett, were first taken up by the Senate and passed by that body with little review. Thus, the House was asked to accept them in the House-Senate Conference without any hearings of their own and without any corresponding House measure.

The House conferees were skeptical--skeptical of NAEP, skeptical of tests, skeptical, especially, of tests for minority students and others in the Chapter I population, because tests did not, per se, improve education for disadvantaged students. The result was that the Senate authorization for a state component was made a "pilot" or "trial" program, and an "independent" evaluation was mandated. The text of the conference report (see Note B) and of the ensuing law (see Note C) are attached. By Washington standards, the conference report was especially detailed (never mind that it is repetitive.) It asked for:

- evaluation of the meaning and reliability of differences in student performance observed across States;

- (amazingly) exploration of "ceiling effects" in high performing States and the link of test content with State curricula; and

- o analysis as to whether NAEP data presentations adequately provide a context for understanding factors that affect education achievement, such as per capita income, per pupil expenditures, ethnic and racial composition and level of urbanization.

The law itself directed the Commissioner of Education Statistics to arrange with "a nationally recognized organization (such as the National Academy of Sciences or the National Academy of Education)" to "assess the feasibility and validity of assessments and the fairness and accuracy of the data they produce. . . describe the technical problems . . and . . . what was learned about how to best report data. . . ."

It was clear that Congress was not about to authorize a permanent State assessment program without further consideration.

NAE was commissioned by the National Center for Education Statistics in October, 1989, to conduct the Trial State Assessment evaluation. The American Institutes for Research, a corporate research organization with long experience in testing, joined the Academy to provide a continuing home for the necessary staff activities.

The Panel is a text-book case of a "blue ribbon" group: The co-chairs are, Bob Glaser and Bob Linn; Executive Director, George Bohrnstedt, AIR. The initial group included Linda Darling-Hammond, Isabel Beck, Lloyd Bond, Ann Brown, Al Shanker, Gordon Ambach, Lyle Jones, David Cohen, Lorrie Shepard, Mike Smith, Ramon Cortines, among others--all well known names in AERA. It functions by commissioning papers, designing empirical studies, some requiring field work as part of the NAEP contract, and extensive discussions among members as to interpretation of the evidence set before them.

To date, 34 studies (see Note D) have been conducted by the Panel on such topics as:

- o Characteristics of the statistical design--sampling, eligibility and exclusion issues, and use of NAEP below the State level;
- o Content and curricular validity of the test;
- o Analysis and reporting--influence of choice of content, statistics and subpopulation breakdowns; validity of the NAEP achievement levels; comparisons of student performance on NAEP with other standardized tests;
- o State and local costs; and
- o Impact of the NAEP Trial State Assessment program.

There have been three major summary reports: (1) on the 1990 Trial, (2) on achievement levels, and (3) on the 1992 Trial and the Panel has made recommendations on a variety of topics (see Note E), for example:

- That State NAEP be continued, but that each grade and each subject be evaluated;
- That NAEP should be more inclusive in its coverage of students with disabilities, LEP students, students in private schools, and out of school youth;
- That management of State NAEP should be modified in several respects, including a tighter requirement for school participation in the sample, permitting annual administration instead of biennial, permitting a reduced sample for small States; and
- And, on the substance and reporting of NAEP, that there be a closer fit of the math test with NCTM standards, that the achievement levels be treated as developmental work and separated from the regular NAEP reports, and international benchmarks be established.

The Academy Panel has briefed Congressional staff on results and sent notices to the press. I have sent reports to Congress and also to all State Departments of Education, as required by law, and, recently, to Governors as well.

INFLUENCE ON NATIONAL POLICY

Congressional action on NAEP's soon to expire legislation has moved only as far as House passage. But here are some observations at this point about the impact of the Panel's work:

The House took action on many issues dealt with in the NAE reports:

- It renewed and upgraded the authorization to continue the State assessments;
- But the State assessments were kept as a "trial" for grades or subjects not already conducted and evaluated;
- The House required that achievement levels be used only on a trial basis until they meet rigorous evaluation criteria established by the Commissioner; reports with achievement levels must be "separate and apart" from regular NCES NAEP reports; and

- The House continued the evaluation requirements and specified that "the National Academy of Education or the National Academy of Sciences" must do them (no longer says "such as. . .") and these are to cover not only State assessments, by also national assessments, LEA assessments, and student performance levels.

I do not mean to make unwarranted claims, here, so these need to be labeled as "interim" conclusions based on the record so far. The Senate has yet to act, and the concluding House-Senate conference has yet to be held.

My conclusions at this interim stage are that the Academy study has been well received and that the intended Congressional audience has paid attention. There is other evidence--including sessions at this year's AERA conference--that additional audiences have paid attention as well.

Why is this the case? I would posit several factors leading to these conclusions are that:

- The Panel covered what Congress asked--not slavishly, but with their own expertise applied;
- They did their homework--formulated studies, carried out empirical work, assessed results, applied judgment;
- They took great pains to comport themselves as impartial judges and were largely successful in that;
- They wrote their results and conclusions in clear English;
- They observed Washington etiquette in submission so NCES sent the report officially to the Hill, as the law specified; and
- They briefed Congress personally.

A COMPARISON WITH AN EARLY MANDATED STUDY

But was this evaluation study unique, a one-time chance, making an impact never before observed and unlikely to recur? Not at all, but some things have changed over a two decade period when Congressionally mandated studies have become increasingly frequent. In preparing for today's session, I recalled another Congressional mandate to evaluate a Federal program, now twenty years ago, and wondered whether any of my observations from this current case study paralleled those made by participants in the earlier one. The evaluation was known as the Compensatory Education Study and it was headed by Paul Hill, then in the National Institute of Education, from 1974 through 1977. He has written about his experience in a 1980 Rand Corporation report entitled "Educational Evaluation in the Public Policy Setting."

Paul described five problems faced by that study and what was done to solve them. Two of the five were (1) guaranteeing that the research was fair and (2) overcoming Congressional distrust of researchers. In the twenty years since authorization of the NIE

study there has been a considerable change in attitude. Congress and the Executive Branch may accuse researchers of irrelevance, sometimes, or of failure to connect their work with real problems. But they do call on researchers; they do mandate evaluations; they do ask researchers to testify; they do ask for briefings on the work; Congressional staff readily meet with researchers and with members of evaluation study panels. Perhaps this is a statement about the education research community as well--a community wanting to make their work count in important places because, finally, those places take actions that can affect American classrooms in powerful ways.

A third problem faced by the Compensatory Education Study Paul Hill described as "resisting pressure from contending parties." In the Compensatory Education case, there were contesting positions among the sponsors of the legislation, so controversy was built into the statute. That was the case, too, with the Trial State Assessment, one dealt with, I believe, by a strong and continuous effort to build balance into the agenda, and through extensive deliberations as to interpretations the Panel would provide. The search for balance goes to great lengths, for example, in the issue of below State use of the NAEP tests--that is, by districts or schools. The Academy conclusion--to strip away the rhetoric--is don't do it, but "If Congress weighs and reads the evidence presented in this report and decides to lift the ban," then only do so at the district level and only with conditions (which are specified).

The fourth problem identified by Paul Hill was moving from broad objectives in the language of the bill to specific researchable problems. In the Trial State Assessment evaluation, the law and the Conference Report included more details about what Congress wanted than in the earlier Compensatory Study. The Academy has had a free hand in formulating its research plan, negotiating with NCES primarily about the level of funding and access to NAEP field work as a source of data. Congress made their primary impact on the study, probably, through their inclusion of language that the study be performed by a group such as the National Academy of Sciences or the National Academy of Education.

And, finally, Paul Hill noted the problem of making results useful to Congress--specifying that deadlines had to be met and that reports must be clear and understandable. Whether they read the record on the Compensatory Education Study experience or not, AIR and the Academy have consistently sought to follow up Paul Hill's advice. Perhaps it has helped to have some members who have much experience in communicating with policy makers.

The summing up is this: Researchers and their work can and do have a strong influence on national policy--

> when they have evidence for their conclusions,
> when their work deals with problems policymakers must solve, and when they can state their conclusions clearly.

Note A

A Nation at Risk: The Imperative for Educational Reform, National Commission on Excellence in Education, Washington, D.C., 1983.

Prince, Cynthia D., Reactions to the Goal 1 Technical Planning Subgroup Report on School Readiness, National Education Goals Panel, Washington, D.C., 1992.

Report of the Goal 2 Technical Planning Subgroup on Core Data Elements/Resolution, National Education Goals Panel, Washington, D.C., 1993.

Assessing Citizenship. The Goal 3 Technical Planning Subgroup on Citizenship Report, National Education Goals Panel, Washington, D.C., 1992.

Assessing National Goals Relating to Postsecondary Education: Goal 5 Task Force Report, National Education Goals Panel, Washington, D.C., 1992.

Report on the Goal 5 Technical Planning Subgroup on International Workforce Skills, National Education Goals Panel, Washington, D.C., 1992.

Wirt, John G., et al., Summary of Findings and Recommendations. National Assessment of Vocational Education Final Report, Volume I, National Assessment of Vocational Education, U.S. Department of Education, Washington, D.C., 1989.

Muraskin, Lana D., The Implementation of the Carl D. Perkins Act. National Assessment of Vocational Education Final Report, Volume II, National Assessment of Vocational Education, U.S. Department of Education, Washington, D.C., 1989.

Reinventing Chapter 1: The Current Chapter 1 Program and New Directions: Final Report of the National Assessment of Chapter 1 Program, Planning and Evaluation Service, U.S. Department of Education, Washington, D.C., 1993.

Alexander, Lamar; James, H. Thomas; and Glaser, Robert, The Nation's Report Card: Improving the Assessment of Student Achievement, Report of the Study Group, National Academy of Education, Cambridge, MA., 1987.

Levine, Daniel B., Ed., Creating a Center for Education Statistics: A Time for Action, National Academy of Sciences, National Research Council, Washington, D.C., 1986.

1988 NAEP Law

(B) then enter into an agreement with the Center for the fiscal year to comply with those information and data gathering requirements.

(3) To establish and maintain the system,[1] the Commissioner—

(A) shall—

(i) provide technical assistance to the States regarding the collection, maintenance, and use of the System's data, including the timely dissemination of such data; and

(ii) to the extent possible, implement standard definitions and data collection procedures; and

(B) may—

(i) directly, or through grants, cooperative agreements, or contracts, conduct research, development, demonstrations, and evaluation activities that are related to the purposes of the System; and

(ii) prescribe appropriate guidelines to ensure that the statistical activities of the States participating in the System produce data that are uniform, timely, and [illegible]

(i)(1) With the advice of the National Assessment Governing Board established by paragraph (5)(a)(i), the Commissioner shall carry out, by grants, contracts, or cooperative agreements with qualified organizations, or consortia thereof, a National Assessment of Educational Progress. The National Assessment of Educational Progress shall be placed in the National Center for Education Statistics and shall report directly to the Commissioner for Educational Statistics. The purpose of the National Assessment is the assessment of the performance of children and adults in the basic skills of reading, mathematics, science, writing, history/geography, and other areas selected by the Board.

(2)(A) The National Assessment shall provide a fair and accurate presentation of educational achievement in skills, abilities, and knowledge in reading, writing, mathematics, science, history/geography, and other areas specified by the Board, and shall use sampling techniques that produce data that are representative on a national and regional basis and on a State basis pursuant to subparagraphs (C)(i) and (C)(ii). In addition, the National Assessment shall—

(I) collect and report data on a periodic basis, at least once every 2 years for reading and mathematics; at least once every 4 years for writing and science; and at least once every 6 years for history/geography and other subject areas selected by the Board;

(ii) collect and report data every 2 years on students at ages 9, 13, and 17 and in grades 4, 8, and 12;

(iii) report achievement data on a basis that ensures valid reliable trend reporting;[3]

(iv) include information on special groups.

(B) In carrying out the provisions of subparagraph (A), the Secretary and the Board appointed under paragraph (5) shall assure that at least 1 of the subject matters in each of the 4 and 6 year cycles

described in subparagraph (A)(i) will be included in each 2 year cycle Assessment.

(C)(i) The National Assessment shall develop a trial mathematics assessment survey instrument for the eighth grade and shall conduct a demonstration of the instrument in 1990 in States which wish to participate, with the purpose of determining whether such an assessment yields valid, reliable State representative data.

(ii) The National Assessment shall conduct a trial mathematics assessment for the fourth and eighth grades in 1992 and, pursuant to subparagraph (6)(D),[1] shall develop a trial reading assessment to be administered in 1992 for the fourth grade in States which wish to participate, with the purpose of determining whether such an assessment yields valid, reliable State representative data.

(iii) The National Assessment shall ensure that a representative sample of students participate in such assessments.

(iv) No State may agree to participate in the demonstration described in this subsection without full knowledge of the process for consensus decisionmaking on objectives to be tested, required in paragraph (6)(E), and of assessment demonstration standards for sampling, test administration, test security, data collection, validation and reporting. States wishing to participate shall sign an agreement developed by the Commissioner. A participating State shall review and give permission for release of results from any test of its students administered as a part of this demonstration prior to the release of such data. Refusal by a State to release its data shall not restrict the reporting of data from other States that have approved the release of such data.

(v) The Commissioner shall provide for an independent evaluation conducted by a nationally recognized organization (such as the National Academy of Sciences or the National Academy of Education) of the pilot programs to assess the feasibility and validity of assessments and the fairness and accuracy of the data they produce. The report shall also describe the technical problems encountered and a description about what was learned about how to best report data from the National Assessment of Educational Progress. The results of this report will be provided to the Congress and to States which participated in assessments pursuant to paragraph (C) (i) and (ii) within 18 months of the time such assessments were conducted.

(D)(i)[2] The National Assessment shall have the authority to develop and conduct, upon the direction of the Board and subject to the availability of appropriations, assessments of adult literacy.

(3)(A) The National Assessment shall not collect any data that are not directly related to the appraisal of educational performance, achievements, and traditional demographic reporting variables, or to the fair and accurate presentation of such information.

(B) The National Assessment shall provide technical assistance to States, localities, and other parties that desire to participate in the assessment to yield additional information described in paragraph (2).

Note B

BEST COPY AVAILABLE

(a) The House recedes.

(b) The Senate amendment, but not the House bill, amends the Carl D. Perkins Vocational Education Act to clarify that single pregnant women including teenagers are eligible for services under the displaced homemaker setaside.

The House recedes with an amendment clarifying that single pregnant women are eligible for participation in vocational education programs.

38. The Senate amendment, but not the House bill, directs the Secretary to cease action regarding the grant procurement process for the National Center for Research in Vocational Education until the GAO has completed a review of this procedure.

The House recedes with an amendment authorizing $2 million for the Ohio State University and $2 million for the University of California at Berkeley to support ongoing activities through December 31, 1988, and providing that such amount shall be deducted from the total award made for a National Center for the 1988 grant award year. This provision is to take effect immediately upon

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

1. The Senate amendment, but not the House bill, amends and expands the scope of the National Assessment for Educational Progress to: report every two years on reading and mathematics, every four years on writing and science, and every six years on history, geography, and civics. In addition, the Senate amendment expands NAEP to provide reports on a national, regional and state basis, and enables the States to participate in the NAEP assessment on a voluntary basis. The authorization is $12.5 million for FY 1989, $18.64 for FY 1990, $17.9 for FY 1991, and $19.6 for each of fiscal years 1992 and 1993.

The House recedes with an amendment limiting the expansion of the National Assessment of Educational Progress (NAEP), both in terms of subject areas and State representative sampling; placing the NAEP in the National Center for Educational Statistics, reporting to the Commissioner of Educational Statistics; requiring a comprehensive study by an independent group of the 1990 and 1992 State representative demonstration assessments; and strengthening the independence of the National Assessment Governing Board.

The Commissioner is authorized to carry out the National Assessment by grants, contracts, or cooperative agreements with qualified organizations or consortia thereof. By this language, the conferees intend that the Commissioner, with the advice of the National Assessment Governing Board, may have either a single grant, contract, or cooperative agreement or any combination of grants, contracts or cooperative agreements.

The expanded National Assessment shall assess the performance of students in reading, mathematics, science, writing, history/geography and other areas selected by the Board. The conferees strongly urge that other areas considered by the Board will include civics and economics.

The expansion of the National Assessment to collect state representative data, on a voluntary basis, is an outgrowth of current

BEST COPY AVAILABLE

1988 NAEP

Conference Report

NAEP practice which lets states obtain such data if they pay the costs of collecting it. The provisions in the Hawkins-Stafford Education Amendments will build on this practice to determine whether an expanded collection of state representative NAEP data is feasible and desirable. The collection of such information will allow participating states to compare themselves to each other and to national NAEP averages. In addition, the data will let state officials monitor their state's progress on NAEP assessments over time.

The conferees wish to emphasize that the purpose of the expansion of NAEP is to provide policy makers with more and better state level information about the educational performance of their school children so that participating states might better measure the educational performance of their children. The goal is not to provide a scorecard by which to rank state educational systems. Data from this assessment is not to be used to compare, rank or evaluate local schools or school districts.

The independent evaluation of the state representative data demonstrations is an important part of the changes authorized here.

The conferees intend the independent evaluation of State representative demonstration assessments will, at a minimum, assess the extent to which differences among States in scores are meaningful and reliable.

The evaluation should also explore the extent to which results are affected by decisions about the test itself. For example, how are the rankings of States altered if the weight given to various skills is changed and what skills should be given priority? Are the rankings of some States affected by the number of high-achieving students whose scores are constrained by ceiling effects?

If possible, the evaluation should also assess the extent to which results are shaped by differences in the closeness of the match between the content of the test and the curricula of the States.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

The conferees intend the independent evaluation of State representative State demonstration assessments will, at a minimum, assess the extent to which differences among States in scores are meaningful and reliable, how well the States participating in the State representative sample do, in fact, constitute a representative sample of the States adequate to assess opportunities and risks in a nation-wide NAEP with state-by-state comparisons. To do so, the evaluation must assess the extent to which results are affected by a variety of extraneous factors. This shall include consideration of the representativeness of participating States in terms of regional representation, ethnic and racial composition, per capita income, curricula, and other variables that influence aggregate indicators of educational achievement, such as which students are excluded from testing (for example, handicapped students or students with limited proficiency in English).

If results are adjusted to take contextual factors into account, the effects of alternative adjustments should be tested. If unadjusted results are presented for groups of States classified in terms of contextual factors, alternative classifications should be assessed.

If possible, the evaluation should also assess the extent to which results are shaped by differences in the closeness of the match between the content of the test and the curricula of the States.

The report shall also assess National Assessment presentations including their effectiveness in providing educators, policy makers, and the general public with useable information and in providing readily understandable information to interpret the strengths and weaknesses of National Assessment findings. The evaluator shall analyze whether National Assessment presentations adequately present data in the context of factors which affect educational achievement including per capita income, per pupil expenditures, ethnic and racial composition and level of urbanization.

The NAEP contractor shall provide the evaluator, in a timely fashion, with the data needed for carrying out the evaluation. For example, data on the characteristics of non-participation will also be provided to the evaluator. The contractor shall also provide the evaluator with information on secure test items needed for analysis. The evaluator, however, will not release secure items to the public, in print or on tape.

(a) The Senate amendment, but not the House bill, changes the name, the membership, and the responsibilities of the current Assessment Policy Committee.

The House recedes with an amendment changing the membership of the National Assessment Governing Board to be more representative of professional educators and testing experts.

Each State choosing to participate in assessments made on a State basis shall cover the cost of coordinating such assessments within the State, in addition to the cost of administering assessments at the school level. Such coordination will include technical assistance to local schools selected for the State sample, securing cooperation of schools, and scheduling tests at times convenient for sample schools, as well as monitoring the sample selection following the design and standards established for State tests.

2. The Senate amendment, but not the House bill, authorizes a new Fund for the Improvement and Reform of Schools and Teaching. This authorization gives the Secretary authority to make grants to SEAs, LEAs, IHEs, and nonprofit organizations to improve the performance of students and teachers. Grants may not be less than $5,000 nor more than $125,000. The Senate amendment further mandates the establishment of a Board to set priorities for awarding grants and to review and evaluate these grants. The authorization is $18 million for FY 1989, $18.9 million for FY 1990, $19.9 million for FY 1991, $21 million for FY 1992, and $22 million for FY 1993.

The House recedes with an amendment to include Pride in Schools as an activity authorized under this part and to add Family-School Partnerships under this part with one-third of the total authorization for this part reserved for this function.

Note C

BEST COPY AVAILABLE

Note D

1

Studies completed by the Evaluation Panel of the National Academy of Education

1992
Assessing Student Achievement in the States
George Bohrnstedt, Project Director

A Critique of Sampling in the 1990 Trial State Assessment
Bruce D. Spencer

Eligibility/Exclusion Issues in the 1990 Trial State Assessment
Bruce D. Spencer

Evaluation of the Implementation of the 1990 Trial State Mathematics Assessment
Donald H. McLaughlin, Francis B. Stancavage, Jay G. Chambers
Elizabeth Hartka, Kadriye Ercikan

The Content and Curricular Validity of the 1990 NAEP Mathematics Items: A Retrospective Analysis
Edward A. Silver, Patricia Ann Kenney, Leslie Salmon-Cox

The Relative Standing of States in the 1990 Trial State Assessment: The Influence of Choice of Content, Statistics, and Subpopulation Breakdowns
Robert L. Linn, Lorrie Shepard, Elizabeth Hartka

A Study of the Impact of Reporting the Results of the 1990 Trial State Assessment: First Report
Frances B. Stancavage, Edward Roeber, George Bohrnstedt

General Issues in Reporting the Results of the NAEP Trial State Assessment
Richard M. Jaeger

The case for District- and School-Level Results from NAEP
Ramsey Selden

Cautions on the Future of NAEP: Arguments Against Using NAEP Test and Data Reporting Below the State Level
Walter Haney, George F. Madaus

Reasonable Inferences for the Trial State NAEP Given the Current Design: Inferences That Can and Cannot Be Made
Edward H. Haertel

1993
Setting Performance Standards for Student Achievement
Lorrie Shepard, Principal Investigator

BEST COPY AVAILABLE

Note D
2

(Background Studies)
An Evaluation of the 1992 NAEP Reading Achievement Levels, Report One: A Commentary on the Process
David Pearson, Lizanne DeStefano

Validity of the 1992 NAEP Achievement-Level-Setting Process
Donald H. McLaughlin

Order of Angoff Ratings in Multiple Simultaneous Standards
Donald H. McLaughlin

Rated Achievement Levels of Completed NAEP Mathematics Booklets
Donald H. McLaughlin

An Evaluation of the 1992 NAEP Achievement Levels, Report Two: An Analysis of the Achievement-Level Descriptors
David Pearson, Lizanne DeStefano

Expert Panel Review of the 1992 NAEP Mathematics Achievement Levels
Edward A. Silver, Patricia Ann Kenny

Comparison of Teachers' and Researchers' Ratings of Students' Performance in Mathematics and Reading with NAEP Measurement of Achievement Levels
Donald H. McLaughlin (and 13 other authors)

Comparisons of Student Performance on NAEP and Other Standardized Tests
Elizabeth Hartka

Comparing the NAEP Trial State Assessment Results with the IAEP International Results
Albert E. Beaton, Eugenio J. Gonzalez

An Evaluation of the 1992 NAEP Reading Achievement Levels, Report Three: Comparison of Cutpoints for the 1992 NAEP Reading Achievement Levels with Those Set by Alternative Means
David Pearson, Lizanne DeStefano

1993
The Trial State Assessment: Prospects and Realities
George Bohrnstedt, Project Director

1994 (Background Studies)
A Study of Eligibility Exclusions and Sampling: 1992 Trial State Assessments
Bruce D. Spencer

A Study of Students Excluded from the 1992 National Assessment of Educational Progress Trial State Assessment
Elizabeth Hartka

BEST COPY AVAILABLE

Note D
3

State and Local Costs of the 1992 Trial State Assessment
Catherine O'Donnell, Jay Chambers, Dey Ehrlich

The Content and Curricular Validity of the 1992 NAEP Reading Framework
Bertram C. Bruce, Jean Osborn, Michelle Commeyras

Evaluation of the 1992 Reading Framework for the National Assessment of Educational Progress
Julia H. Mitchell

The Content and Curricular Validity of the 1992 NAEP TSA in Mathematics
Edward A. Silver, Patricia Ann Kenney

Content Validation of the 1992 NAEP in Reading: Classifying Items According to the Reading Framework
David Pearson, Lizanne DeStefano

Impact of the 1992 NAEP Trial State Assessment Program: A Followup Study
Francis B. Stancavage, Edward D. Roeber, George W. Bohrnstedt

Issues in the Development of Spanish-Language Versions of the National Assessment of Educational Progress
Walter G. Secada

The Judged Congruence Between Various State Assessment Tests in Mathematics and the 1990 National Assessment of Educational Progress Item Pool for Grade-8 Mathematics
Lloyd Bond, Richard M. Jaeger, assisted by Sarah E. Putnam

A Study of the Administration of the 1992 National Assessment of Educational Progress Trial State Assessment
Elizabeth Hartka, Donald H. McLaughlin

Studies Proposed by the National Academy of Education for 1994-95

NOTE: The list of titles and authors for these proposed studies is tentative.

1) **Study of IEP and LEP Exclusions**
George Bohrnstedt, Frances Stancavage

2) **Alternative Assessments for IEP and LEP students**
George Bohrnstedt, Frances Stancavage

BEST COPY AVAILABLE

Note D-4

3) **Study of the 1994 Exclusions and Sampling Frame**
Donald H. McLaughlin

4) **Study of the 1994 TSA Administration**
Elizabeth Hartka

5) **Impact of Public School Nonparticipation**
Donald H. McLaughlin

6) **Combining State and National NAEP**
Edward Haertel

7) **Linking State and National Assessments**
Richard Jaeger

8) **Anchoring Achievement Levels**
Donald H. McLaughlin

9) **Content Validity of the 1994 Reading Assessment**
David Pearson

10) **Impact and Reporting of NAEP Results**
Edward Roeber, Frances Stancavage

11) **Study of NAEP Scaling: Trends, Content, Mode of Assessment**
Robert Linn, Donald McLaughlin

12) **Acquisition of Competence and its Relevance for NAEP Assessments**
Robert Glaser, Donald McLaughlin

13) **The Capstone Report**
Robert Glaser, Robert Linn, George Bohrnstedt

BEST COPY AVAILABLE

Note E

Topics of NAE Recommendations

Continuation of State NAEP
- Continue State NAEP but conduct evaluations of grades and subjects not covered and merge State evaluation with national

Coverage of State NAEP
- students in private schools
- IEP (study exclusions)
- Out of school 17 yr. olds OR cohort in NALS
- Trial in Spanish for LEP students
- Not below State level OR very limited and not to schools

Management of State NAEP
- Continue to monitor sites for uniformity
- Tighter requirement for school participation rates
- Merge State and national samples for efficiency
- Permit annual administration to distribute workload
- Use half-sample size for small States
- Do not increase State cost sharing
- Provide adequate funding for a quality program

Substance and reporting of the Assessment
- Text coverage and item types more consistent with NCTM
- Set achievement levels in coordination with NESIC
- Use focus groups to help determine useful displays
- Provide examples for press of proper data interpretations
- Discontinue use of Angoff methods for achievement levels
- Discontinue reporting by achievement levels as used in 1992
- Ask for standard setting advice from more diverse sources (such as business leaders, standards committees, content experts)
- Publish achievement levels separate from official reports
- Use percentile scores to monitor achievement
- Use international comparisons to set benchmarks and provide for equating with TIMSS
- Work with NEGP to determine how to report over the 1990s
- Implement within-grade score reporting

Long term recommendations on Performance Standards
- Develop content standards and performance standards in an iterative process
- Continuous oversight group from frameworks through reporting
- Address issue of developmental model that underlies achievement levels and scales
- Evaluate achievement levels before use for regular reporting
- Recognize need for multiyear developmental process
- Provide for stability of measures over an 8 to 10 year period